Data Warehouse Conceptual Design-A Literature Survey

[1]Sharmila Mat Yusof, [2]Fatimah Sidi, [2]Hamidah Ibrahim and [2]Lilly Suriani Affendey
[1]School of Computing, College of Arts and Sciences, Universiti Utara Malaysia,
06010 Sintok, Kedah, Malaysia
[2]Faculty of Computer Science and Information Technology, Universiti Putra Malaysia,
43400 Serdang, Selangor, Malaysia

Abstract: Data Warehouse (DW) can provide an excellent approach in transforming Online Transaction Processing (OLTP) data into useful and reliable information to support organization's decision making. As such, it can be a basis for data analysis techniques such as multidimensional analysis and data mining. However, DW design process has been known as a complex task that requires systematic and structured approach to guarantee its success. Therefore, there have been various methodologies proposed to carry out the design process which can be classified into requirement-driven, data-driven and hybrid approaches. In this study, a literature survey was made to obtain related works by a set of pertinent key words related to DW conceptual design. The objective of this survey is to provide the state of the art of DW conceptual design methodologies in narrative and summarized forms. The main contribution is to provide understanding of the trend, issues and solutions proposed to date in DW conceptual design and along the way to discover the novel and great contribution works that form an important basis in the DW conceptual design.

Key words: Data warehouse, conceptual design, multidimensional model, surve, provide, DW conceptual

INTRODUCTION

Data Warehouse (DW) provides an excellent approach in transforming Online Transaction Processing (OLTP) data into useful and reliable information to support organization's decision making. As such, it provides the basis for data analysis techniques such as Multi Dimensional (MD) analysis and data mining. However, building a data warehouse requires adopting design and implementation techniques completely different from those underlying OLTP systems (Golafarelli *et al.*, 1998; Cravero and Sepulveda, 2012). Building DW requires comprehensive design method that guarantees the reveal of strategic hidden values in various operational data sources and at the same time meeting organizational business goals. DW design method follows the typical of database system which start with the conceptual design, logical and physical design (Rizzi *et al.*, 2006). Conceptual design, being the first design step is the most crucial step because it is an essential ground work for building a DW which is well-documented and completely satisfies business requirements (Romero and Abello, 2009). Thus, we only focus on the conceptual design part of the DW here.

A literature survey was made to obtain works related to DW conceptual design by a set of pertinent keywords (i.e., data warehouse, conceptual design and multidimensional model). The articles were examined and those focusing on DW conceptual design are chosen. In addition, several survey articles (Cravero and Sepulveda, 2012; Romero and Abello, 2009; Jindal and Taneja, 2012; Elhaj and Jamel, 2014; Cravero and Samuel, 2014) have also been examined to trace previous articles that have been studied and have been included in this survey in order to discover the origin and significant works that contribute to DW conceptual design. Works that are not focusing on the conceptual design part were omitted. In case for similar works by the same researchers, the latest one is considered. Thus, this survey intent to contribute on the understanding of the trend, issues and solutions proposed to date in DW conceptual design and along the way to discover the novel and great contribution works that form an important basis in the conceptual design.

The main objective is to provide a general overview to those novices interested in exploring the progress of DW conceptual model thus far. With the presentation of the works in narrative and summarized form, it is hoped that this study can provide a general overview of the

---

Corresponding Author: Sharmila Mat Yusof, School of Computing, College of Arts and Sciences, Universiti Utara Malaysia,
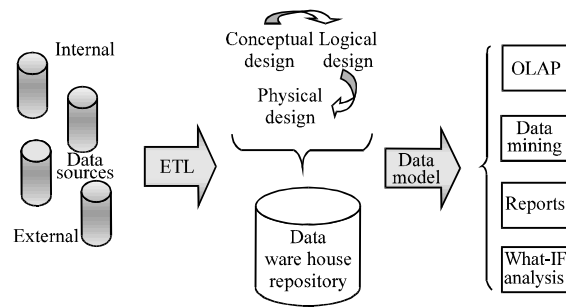06010 Sintok, Kedah, Malaysia

Fig. 1: Data warehouse design process

current state of the art of DW conceptual design and its trend, issues and proposed solution. The summaries with specific criteria are presented in study. Those interested in details of the approach can refer the original articles as in references.

Background of data warehouse: The classical definition of DW was coined by Inmon (1996) as being a collection of historical, subject-orientated, non-volatile, integrated data which has been designed to support an organization's decision-making.

A DW may generally be classified into three main phases which starting from Extraction, Transformation and uploading of data (ETL) from different data sources into a DW, next, data modeling design (i.e., conceptual design) and finally efficient and flexible access to the data. In the second phase, the extracted data must be presented in such a way that easily accessible by analytical tool such as Online Analytical Processing (OLAP). This presentation is done by way of design step which start with conceptual design, logical then only the physical itself. In conceptual design phase which is the focus of the survey, its output is the technology-independent conceptual model that contributes to the basis of the effective representation of data in DW. Figure 1 illustrates the three main phases of the DW design process.

In DW conceptual design, it is widely acceptable to follow the MD model due to its features that best suited for analysis tools such as OLAP (Kimball and Ross, 2009). This is because OLAP tools can best extract relevant data from DW based on its multidimensionality (i.e., fact, measure and dimension). From the survey, MD modeling approaches can be generally categorized into three main methodologies which are requirement driven, data driven and the hybrid which is the combination of the previous two. As for the hybrid methodology, it can be classified into two categories which are pure hybrid methodology and integration-derived methodology which then further classified into sequential and parallel (Romero and Abello, 2009; Winter and Strauch, 2004; Tria *et al.*, 2012) (Fig. 1).

MATERIALS AND METHODS

Conceptual design methodologies: In DW conceptual design, the initial common approaches were on the requirement-driven (or demand-driven) and data-driven (or supply driven) and later the hybrid approach was introduced to take advantages on both approaches. In requirement-driven approach, the conceptual design is based on the information needs of decision makers while data sources are considered separately during ETL activities. On the other hand, data-driven approach is based on a detailed analysis of data sources and information needs of decision makers are considered later after the implementation of the conceptual design. A hybrid approach seeks to combine both the requirement-driven and data-driven to design a DW from data sources but taking end-user's needs into account.

Requirement-driven: Kimball (1996) propose a requirement-driven approach by giving informal guideline in identifying MD elements based on example. Kimball approach to DW design is actually from the building of the data marts which interest to particular organization. Thus, the approach starts with the identification of data marts, preferably by looking at the data sources. Then, for each data marts, the dimensions associated are determined by forming a matrix table to reveal the shared dimensions. Then finally fact table will be designed by following the three-step rules which consists of determining the grain of detail, choosing the dimensions for the fact table and adding measures to the fact.

Winter and Strauch (2004) introduce a requirement-driven approach by giving a comprehensive coverage to the information requirement analysis based on case studies and literature review. The manual incremental approach comprises of four phases of initialization, "As is" analysis, "To be" analysis and modeling phase. The conciliation of the requirements and data sources is done during "To be" analysis phase.

Prat *et al.* (2006) present a requirement-driven approach focusing on the transformation from conceptual to logical and physical design of DW by using Unified Modeling Language (UML) standard. They used interviews, joint session and report analysis to elicit user requirements but no formal guideline was discussed.

Prakash and Gosain (2008) suggest a requirement driven approach where they use organizational goals to identify DW information content. From the organizational goals, the Goal-Decision-Information (GDI) schema is

Table 1: Requirement-driven approach

Researchers	REF	RP/T	RF
Kimball (1996)	Informal	End user (business requirement definition)/interview	Fact schema
Winter and Strauch (2004)	Informal	End user (business question)/case study (questionnaire, interview)	Not clearly specified
Prat *et al.* (2006)	Informal	Decision maker (informal statement)/interview, joint session and report analysis	UML class diagram
Prakash and Gosain (2008)	Formal	Organizational perspective (Organizational goals)/Goals and scenarios modeling	Goal-Decision Information (GDI) schema converted to SSQL
Sarkar (2012)	Formal	Business process/not clearly specified	Graph Object Oriented MD (GOOMD) model

created from the set of goals, decisions and information. The information requirements to support decisions are then formulated into the technical form using Specification SQL (SSQL).

Sarkar (2012) presents a requirement-driven approach by developing a Business Object (BO) based requirement analysis framework which consists of three main steps namely, early requirement analysis, detailed requirement analysis and mapping. During mapping, the requirement descriptions are mapped with the data model to ensure data source availability. The proposed framework supports abstraction mechanism and reuse of the elements that have been used to realize the different business concepts of the domain and useful for analytical task. These frameworks feature suits well with the current nature of the evolving and dynamic business processes.

Table 1 summarizes the requirement-driven approach according to Requirement Elicitation Formalism (REF), Requirement Perspective and Technique (RP/T) and requirement Representation Format (RF).

Data-driven: Golfarelli *et al.* (1998) propose a data-driven approach that focus on the analysis of the data sources to come out with the conceptual design which is based on Dimensional Fact Model (DFM). The DFM is a model that builds cubes and dimensions starting from the relational or Entity Relationship (ER) schemas of the data sources. In this approach, the methodology to carry out the conceptual modeling is semi-automatic in which it is based on an algorithm that generates an attribute tree representing an integrated data source. The root of the tree is the relation (entity) arbitrarily chosen by the designer as fact. Then, the algorithm generates a node for each attribute encountered while navigating recursively through the relations (entities) via. the joins paths (relationships). Later, designer can edit the tree by adding and removing new nodes or change the parent nodes.

Phipps and Davis (2002) introduce data-driven approach that automatically generates conceptual schemas using Multidimensional ER (M E/R) model that consider input from OLTP relational schema. They introduce an algorithm to automatically create the whole candidate conceptual schemas from OLTP data sources. After all the candidate schemas have been created, the evaluation algorithm is used to refine the derived schema to ensure user requirements are met.

Moody and Kortink (2003) present an approach for developing MD schemas via. an ER model. Although, they were not the first to use an ER Model, they presented a structured, formal methodology for developing conceptual schema. However, the approach was done manually.

Jensen *et al.* (2004) propose a semi-automatic data-driven approach that use data mining technique to identify the MD elements in relational model automatically. However, the fact elements need to be manually identified by users. Thus, this approach is considered as semi-automatic. The final output is the snowflake schema derived from the hierarchies of dimensions discovered.

Romero and Abello (2010a, b) have expanded their earlier work of hybrid approach by focusing on data-driven approach. The analyses of data sources are used to guide the elicitation of the end-user requirements (in the form of user queries). Their user-centered fully automatic data-driven approach named as AMDO (Automating Multidimensional Design from Ontologies) consists of three main steps starting from data sources analysis, multidimensional concept identification and automatic schema generation.

Hachaichi and Feki propose an automatic data-driven method starting from object oriented database. The method consists of three phases which are schema retrieval, MD concepts construction and schema refinement. They present an algorithm to extract relevant MD concepts to produce the relevant DFM.

Table 2 summarizes the data-driven approach according to Automation Level (AL) Manual (M)/Semi-Automatic (SA)/Automatic (A), Input Abstraction (IA), Output Abstraction (OA) and Schema Evaluation (SE).

Hybrid: Bonifati *et al.* (2001) present a sequential hybrid approach that consists of three main steps starting from user requirement collection, detailed analysis of data sources and finally schema integration. In

Table 2: Data-driven approach

Researchers	AL	IA	OA	SE
Golfarelli *et al.* (1998)	SA	ER schema (conceptual)	DFM (conceptual)	None
Phipps and Davis (2002)	A	Logical(relational schema)	Conceptual(ME/R schema)	TPC-H benchmark
Moody and Kortink (2003)	M	Conceptual (ER schema)	Logical (Star schema)	ER schema illustration by example
Jenson *et al.* (2004)	SA	Logical (relational schema)	Logical (snowflake schema)	None
Romero and Abello (2010a, b)	A	Conceptual (Ontology)	Logical (Constellation star schema)	Car rental case study
Hachaichi	SA	Logical (object data source)	Dimensional Fact Model (DFM)	Prototypical tool (CAME-BDO)

user requirement collection activities, interviews are done to collect information regarding organizational goals and needs from respective stakeholders using a set of forms from Goal/Questions/Metric (GQM) model (Basili *et al.*, 1994). In data source analysis step, the ER schema is converted into a star join graph automatically. Finally, the two schemas from both steps earlier are integrated and ranked in semi-automatic way.

Giorgini *et al.* (2008) propose a parallel hybrid approach which is based on the Tropos methodology (Bresciani *et al.*, 2004) that takes two perspectives of stakeholders' goals (organizational modeling) to identify functional requirements and the decision maker's goals (decisional modeling) to identify the non-functional ones. In this approach, the information requirements are collected using interview where the related diagrams were generated from those requirement findings. The diagrams produced are then used to map with the data sources following three steps of requirement mapping, hierarchy construction where fact schema are generated and finally the schema is refined to fully meet user's expectations.

Mazon *et al.* (2007) present a sequential hybrid approach which start with requirement analysis using i* framework (Yu, 1995). The i* framework is adapted to model the organizational goals and information requirements for DW. The conceptual schema from requirements (expressed in UML standard) is then automatically verified with data source's relational schema using Query/View/Transformation (QVT) relations (Mazon *et al.*, 2006). Their novelty contribution is the systematic, well-structured and comprehensive way employment of hybrid approach in designing the MD model which can be effectively and automatically transform to its corresponding logical model.

Romero and Abello (2010a, b) propose a pure hybrid approach where they consider user-requirements and data sources simultaneously. This methodology suits the scenario where the user requirements are known at the time of the design. In this study, the main aim is to automate most of the design tasks in a hybrid manner and deals with relational data sources. In this approach, user requirements are formalized into SQL queries by users. Then, the SQL queries will be validated using Multidimensional Design by Examples (MDBE)

whether each of them can produce a meaningful MD cube and a graph of a MD model is created.

Francesco present a sequential hybrid approach that covers from requirement analysis until the data modeling part. The proposed methodology introduces the two additional steps of attribute tree generation which can be done automatically and advanced data modeling which is done manually. This approach uses the i* framework (Yu, 1995) to capture user requirements and extended DFM1 for data source view in the reconciliation process. In their approach, the formalization of user requirements is represented using UML. Due to the high level of standardization and formal representation of UML MD schema, the design can be effectively automated.

Thenmozhi and Vivekanandan (2012) come out with a hybrid approach that concerns the semantic issues that arise due to the heterogeneity of the data sources. Their approach consists of a comprehensive framework that produce MD schema from several ontology sources based on requirements. In this approach, user requirements and data sources are conciliated at early design stage. They use a set of ontology matching algorithm in the conciliation process. Then, reasoning algorithms are used to automatically produce the facts and dimensions. Their approach is different from other ontology multidimensional design approach in terms of user requirements can be represented using natural language, interesting MD concepts are identified at early design stage and its derivation is automatically done using reasoning.

Elhaj and Jamel (2014) suggest a parallel hybrid which combines business user requirements and structured data sources in semi-automatic way emphasizing on the involvement of DW designer throughout the process. They also highlighted the importance of ontology usage in order to overcome the semantic and heterogeneity issues. However, they propose to use general semantic resource of Word Net as they claim could reduce time and cost of the design. However, the proposed method is yet to be tested using prototype.

Table 3 summarizes the hybrid approach according to Hybrid Type (HT) Pure (P)/Parallel (PL)/Sequential (S), Automation Level (AL)-Manual (M)/Semi-Automatic

Table 3: Hybrid approach

Reseachers	HT	AL	RT/R	DT/R	FS
Bonifati *et al.* (2001)	S	SA	Interview (GQM model)/star schema	Reengineering process/ER schema	Star schema
Giorgini *et al.* (2008)	PL	SA	Interview (Tropos methodology)/ organizational and decisional model	Reengineering process/relational schema	Fact schema
Mazon *et al.* (2007)	S	SA	Interviews (i* framework)/MD UML class diagram	Reengineering process/relational schema	MD class diagram
Romero and Abello (2010)	P	A	initial work load/SQL queries	Reengineering process/relational schema	Constellation schema
Francesco	S	SA	Interview (i* framework)/MD UML schema	Reengineering process/DFM	Fact schemas
Thenmozhi and Vivekanandan (2012)	P	A	Interview/SKOS format	Reengineering process/ontology schema	Facts and dimensions
Elhaj and Jamel (2014)	PL	SA	Not specified/star schema	Reverse engineering/ER tables	Star schemas

(SA)/Automatic (A), Requirement-driven Technique and Representation (RT/R), Data-driven Technique and Representation (DT/R) and Final Schema (FS).

RESULTS AND DISCUSSION

Discussion on trends and issues: In requirement-driven methodologies, the research by Kimball (1996) and Yu (1995) (in hybrid approach) were mostly referred to and seem to contribute much to the DW conceptual design. While Kimball (1996) provides rather an informal guideline in construction of dimensional model from user requirements on the other hand, Yu (1995) provides us with the set of formal and structured method to capture user requirements from business goals. The trend of capturing user requirements is shifting from informal end-user elicitation tasks to formal and structured stakeholders that play significant role in business process. This trend may mainly due to the changing role of technology from business automation support to business competitive advantage. From the survey in strict requirement-driven approach, requirement collections are mostly done manually by designers which depend on their experience and capability. Therefore, the output of the design is heavily affected by their expertise. Thus, the issue arise from this approach is how to release the large dependency on designer. Thus, the latter efforts have been towards automating of the user-requirement collection activities. However, in order to accomplish this, the requirements need to be formalized as the existing requirements are commonly expressed in natural language that is hardly automatable. Therefore, works such as (Prakash and Gosain, 2008; Romero and Abello, 2010a, b) have proposed to formalize such requirements in formal form such as SQL queries. Another, issue arise from requirement-driven approach is that users may not realize the potential analysis data hidden in heterogeneous data sources as they are not considered in requirements collection activities. Therefore, from the survey, the trend of requirement-driven approach is the

incorporation of data sources conciliation additional step such as works by Winter and Strauch (2004) and Sarkar (2012). Nevertheless, their approach cannot be considered as hybrid because the conciliation on data sources is not given much emphasize.

In data-driven methodologies, the detail analysis of operational data sources is done either using ER conceptual schema or relational logical schemas depend on the availability of the data sources schemas level. From literature, most of the data-driven approach designs are done either manually (Moody and Kortink, 2003), semi-automatic (Golfarelli *et al.*, 1998); Jensen *et al.*, 2004) and fully-automatic way (Phipps and Davis, 2002; Romero and Abello, 2010a, b) (in hybrid approach). Thus, the trend in data-driven approach is shifting from manual to fully automated of design activities. In manual data-driven approach, the extensive analysis is putting a heavy burden to designers when the sizes of the data sources involved are huge. In addition, it also needs to depend on designer's expertise which results in cases similar to requirement-driven approach where sometimes the hidden analysis potentials may be missed. This issues of the burden and dependency on designer's expertise in data-driven approach feasibly can be solved with fully automating the design tasks. Thus, from the survey, ontology is proposed to provide a means to automate the MD design of DW and help to populate data in a more significant manner. Upon this realization, ontology has been incorporated in the works by Elhaj and Jamel (2014), Romero and Abello (2010a, b), Thenmozhi and Vivekanandan, 2012). The adoption of ontology could also solve the syntactic and semantics inconsistency that arise due to the heterogeneity of data sources. Elhaj and Jamel (2014), researchers claim that their approach of using general ontology (i.e., Word Net) could reduce the time and cost of the design. However, there is no further validation on the claim yet. Meanwhile, in full automatic data-driven design, another issue arises which is the generation of too many irrelevant results that are not in the interest of the organizational business. Thus, again

Thenmozhi and Vivekanandan (2012) have introduced filtering function and searching patterns in their approach to cater this problem.

In order to resolve drawbacks of the previous two approaches, current effort on hybrid approach has concentrated on incorporating both approaches together to produce better design outcome that meets both user requirements and data availability. Among the hybrid approaches, the requirement-driven part and data sources analysis part can be considered simultaneously (pure hybrid) in parallel or sequentially. In hybrid approaches, the trend is to provide structured and formal method of user requirement collection and fully automatic data sources analysis. From survey, it can be seen that the requirement-driven part is generally adopting business goals oriented similar to i* framework proposed by Yu (1995). While, in data-driven part, the DFM method (Golfarelli *et al.*, 1998) which consist of attribute tree generation has been applied in most hybrid approaches. Issues that researchers can see in hybrid approach is that whether the fully automation of hybrid approach is actually can provide the optimum solution among all. This is because if we examine the works presented chronologically while most researchers propose to fully automate the design tasks, recent work by Elhaj and Jamel (2014) has emphasized back the manual refinement of schemas (i.e., designer's intervention) to guarantee the optimum solution achieved. In general from the survey, we can see that current works are proposing the hybrid approaches with the improvement effort on both requirement-driven and data-driven part. As the trend is towards automation of most of design tasks, the activities that are most being concentrated in current approach are the formalization of user requirement collection task such as use of UML by Francesco and effective searching method in identifying multidimensional elements in data-driven task such as effort by Thenmozhi and Vivekanandan, (2012).

CONCLUSION

This survey has presented a classification of conceptual design methodologies that have been practiced for DW development which comprised of requirement-driven, data-driven and hybrid approaches. Further under hybrid approach, the alternatives are the sequential or parallel method.

This survey has provided a general overview of the current state of the conceptual DW design in narrative nature and summarization. The trend in the design seems to adopt hybrid methodology with the emphasis on achieving automation on the most part of the design. The purpose is to relieve the burden on having to heavily rely on the designer's expertise, to avoid time consuming task of huge data sources analysis and avoid overlooking potential analysis data in achieving the optimum output. In accomplishing the automation, the formalization of user requirements and use of ontology is seemed to be necessary.

In addition, from the literature survey carried out, the works from (Golfarelli *et al.*, 1996; Kimball, 1996; Yu, 1995) have been a huge contribution as a basis approach in other methodologies proposed from the beginning. Future work will be on detail comparative study of the hybrid methodologies that can provide an optimum solution to support decision making and propose a framework to access its performance.

REFERENCES

Basili, V.R., G. Caldiera and H.D. Rombach, 1994. The goal question metric approach. Encycl. Software Eng., 2: 528-532.

Bonifati, A., F. Cattaneo, S. Ceri, A. Fuggetta and S. Paraboschi, 2001. Designing data marts for data warehouses. ACM. Trans. Software Eng. Method., 10: 452-483.

Bresciani, P., A. Perini, P. Giorgini, F. Giunchiglia and J. Mylopoulos, 2004. Tropos: An agent-oriented software development methodology. Autonomous Agents Multi-Agent Syst., 8: 203-236.

Cravero, A. and S. Samuel, 2014. Multidimensional design paradigms for data warehouses: A systematic mapping study. J. Software Eng. Appl., 7: 53-61.

Cravero, A. and S.S. Sepulveda, 2012. A chronological study of paradigms for data warehouse design. Eng. Res., 32: 58-62.

Elhaj, E. and F. Jamel, 2014. Toward an ontology based approach for data warehousing: State of the art and proposal. Proceedings of the International Conference on Arab Information Technology (ACIT2014), December 09-11, 2014, University of Nizwa, Nizwa, Oman, pp: 170-179.

Giorgini, P., S. Rizzi and M. Garzetti, 2008. GRAnD: A goal-oriented approach to requirement analysis in data warehouses. Decis. Support Syst., 45: 4-21.

Golfarelli, M., D. Maio and S. Rizzi, 1998. The dimensional fact model: A conceptual model for data warehouses. Intl. J. Cooperative Inf. Syst., 7: 215-247.

Inmon, W.H., 1996. Building the Data Warehouse. 2nd Edn., John Wiley and Sons, Canada, ISBN: 0471-14161-5.

Jensen, M.R., T. Holmgren and T.B. Pedersen, 2004. Discovering multidimensional structure in relational data. Proceedings of the 6th International Conference on Data Warehouse Knowledge Discovery, September 1-3, 2004, Springer, Zaragoza, Spain, pp: 138-148.

Jindal, R. and S. Taneja, 2012. Comparative study of data warehouse design approaches: A survey. Intl. J. Database Manage. Syst., 4: 33-45.

Kimball, R. and M. Ross, 2009. The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. 2nd Edn., John Wiley & Sons, New York, USA., ISBN:0-471-20024-7, Pages: 419.

Kimball, R., 1996. The Data Warehouse Toolkit: Practical Techniques for Building Dimensional Data Warehouses. 1st Edn., John Wiley and Sons, New Jersey.

Mazon, J.N., J. Trujillo and J. Lechtenborger, 2006. A set of QVT relations to assure the correctness of data warehouses by using multidimensional normal forms. Proceedings of the 25th International Conference on Conceptual Modeling (ER), November 6-9, 2006, Springer, Tucson, Arizona, USA., pp: 385-398.

Mazon, J.N., J. Trujillo and J. Lechtenborger, 2007. Reconciling requirement-driven data warehouses with data sources via. multidimensional forms. Data Knowl. Eng., 23: 725-751.

Moody, D.L. and M.A.R. Kortink, 2003. From ER models to dimensional models: Bridging the gap between OLTP and OLAP design. J. Bus. Intell., 8: 7-24.

Phipps, C. and K.C. Davis, 2002. Automating data warehouse conceptual schema design and evaluation. Proceedings of the 4th International Workshop on Design and Management of Data Warehouses, May 2•7, 2002, CEUR-WS Publisher, Toronto, Canada, pp: 23-32.

Prakash, N. and A. Gosain, 2008. An approach to engineering the requirements of data warehouses. Requirements Eng., 13: 49-72.

Prat, N., J. Akoka and I. Comyn-Wattiau, 2006. A UML-based data warehouse design method. Dec. Support Syst., 42: 1449-1473.

Rizzi, S., A. Abello, J. Lechtenborger and J. Trujillo, 2006. Research in data warehouse modeling and design: Dead or alive?. Proceedings of the 9th ACM International Workshop on Data Warehousing and OLAP, November 10, 2006, ACM, Arlington, Virginia, ISBN:1-59593-530-4, pp: 3-10.

Romero, O. and A. Abello, 2009. A survey of multidimensional modeling methodologies. Intl. J. Data Warehouse Min., 5: 1-23.

Romero, O. and A. Abello, 2010b. A framework for multidimensional design of data warehouses from ontologies. Data Knowl. Eng., 69: 1138-1157.

Romero, O. and A. Abello, 2010a. Automatic validation of requirements to support multidimensional design. Data Knowl. Eng., 69: 917-942.

Sarkar, A., 2012. Data warehouse requirements analysis framework: Business-object based approach. Intl. J. Adv. Comput. Sci. Appl., 3: 25-34.

Thenmozhi, M. and K. Vivekanandan, 2012. An ontology based hybrid approach to derive multidimensional schema for data warehouse. Intl. J. Comput. Appl., 54: 36-42.

Tria, F.D., E. Lefons and F. Tangorra, 2012. Hybrid methodology for data warehouse conceptual design by UML schemas. Inf. Software Technol., 54: 360-379.

Winter, R. and B. Strauch, 2004. Information requirements engineering for data warehouse systems. Proceedings of the ACM Symposium on Applied Computing, March 14-17, 2004, ACM, Nicosia, Cyprus, ISBN:1-58113-812-1, pp: 1359-1365.

Yu, E., 1995. Modelling strategic relationships for process reengineering. Ph.D Thesis, University of Toronto, Toronto, Ontario.